POLYNOMIAL TERM STRUCTURE MODELS

SI CHENG AND MICHAEL R. TEHRANCHI
UNIVERSITY OF CAMBRIDGE

Abstract. In this article, we explore a class of tractable interest rate models that have the property that the price of a zero-coupon bond can be expressed as a polynomial of a state diffusion process. Our results include a classification of all such time-homogeneous single-factor models in the spirit of Filipovic's maximal degree theorem for exponential polynomial models, as well as an explicit characterisation of the set of feasible parameters in the case when the factor process is bounded. Extensions to time-inhomogeneous and multi-factor polynomial models are also considered.

1. Introduction

A time-homogeneous factor model of the risk-free interest rate term structure is one in which the time-t spot interest rate is of the form

$$r_t = R(Z_t)$$

and the time-t price of a zero-coupon bond of maturity T is of the form

$$P_t(T) = H(T-t, Z_t)$$

where $R : I \to \mathbb{R}$ and $H : \mathbb{R}_+ \times I \to \mathbb{R}$ are given non-random functions and $Z = (Z_t)_{t\geq 0}$ is a given time-homogeneous Markov process with state space I modelling some underlying economic factor. For such a model to be sensible, the functions R and H and the process Z must be intimately connected. The goal if this paper is to study this connection when the function $H(T-t,\cdot)$ is assumed to be a polynomial.

The motivation for this study is classical. Recall that a sufficient condition for the bond market to have no arbitrage is that there exists an equivalent probability measure $\mathbb{Q}$ under which the discounted bond prices $\tilde{P}(T)$, defined by

$$\tilde{P}_t(T) = e^{-\int_0^t r_s ds} P_t(T),$$

are local martingales for all $T \geq 0$. Furthermore, a sufficient condition to ensure that the discounted bond prices are local martingales is the union of the following two assumptions: firstly, an analytic assumption that the function H is a classical solution of the partial differential equation

$$\partial_x H = \sum_{1\leq i\leq d} b_i\, \partial_{z_i} H + \frac{1}{2} \sum_{1\leq i,j\leq d} a_{ij}\, \partial_{z_i z_j} H - RH \text{ on } \mathbb{R}_+ \times I \tag{1}$$

Date: September 2, 2016.
Keywords and phrases: term structure, interest rates, polynomial models.
Mathematics Subject Classification 2010: 91G30, 91B25, 91G80.

for a given spacial domain $I \subseteq \mathbb{R}^d$ and for given functions $b : I \to \mathbb{R}^d$ and $a : I \to \mathbb{R}^{d\times d}$, with boundary condition

$$H(0, z) = 1 \text{ for all } z \in I; \tag{2}$$

and secondly, a probabilistic assumption that Z is a weak solution of the stochastic differential equation

$$dZ_t = b(Z_t)dt + \sigma(Z_t)dW_t, \tag{3}$$

taking values in I for all $t \geq 0$, where W is a Brownian motion under the fixed pricing measure $\mathbb{Q}$ and $a = \sigma\sigma^\top$.

In principle, the above partial differential equation (1) with boundary condition (2) can be solved numerically whenever the functions b, σ and R are suitably well-behaved. However, resorting to a numerical method to solve the partial differential equation can be too slow to use practice for the purpose of calibrating the parameters of the model, and indeed, it obscures the relationship between the dynamics of the factor process and the resulting bond prices. Therefore, there has been considerable interest in developing tractable models, where the function H is of reasonably explicit form.

Perhaps the two most famous tractable factor models are those of Vasicek [17] and Cox, Ingersoll & Ross [4]. In these models the factor is scalar and identified with the spot interest rate, so in the notation above, $d = 1$ and $R(z) = z$, while the functions b and a are affine and the function H is of the exponential affine form

$$H(x, z) = e^{h_0(x)+h_1(x)z}.$$

It is easy to see that the partial differential equation (1) reduces to a system of coupled Riccati ordinary differential equations for the functions h_0 and h_1 and the boundary condition (2) becomes $h_0(0) = h_1(0) = 0$. Furthermore, it is well-known that in both cases the corresponding stochastic differential equation (3) always has a unique *local* solution. While the local solution to the Vasicek stochastic differential equation is in fact the unique global solution, the situation with the Cox–Ingersoll–Ross stochastic differential is more delicate: for some values of the parameters, local solutions may explode in finite time by hitting the boundary of the state space. Duffie & Kan [7] studied exponential affine models where the factor process is of arbitrary dimension $d \geq 1$, finding conditions under which the corresponding stochastic differential equation (3) has a non-explosive solution. Subsequently, there has been a considerable body of research on the properties of these exponential affine models. A notable contribution to this literature is a general characterisation of exponential affine term structure models by Duffie, Filipović & Schachermayer [6].

An exponential affine model can be considered a special case of the family of exponential quadratic models. An early example of a quadratic model was proposed by Longstaff [15], and has since been developed and generalised by Jamshidian [12], Leippold & Wu [14], and Chen, Filipović & Poor [2] among others.

One may wonder if there exist non-trivial exponential cubic (or higher degree) models. Filipović answered this question in the negative, by showing that the maximal degree for exponential polynomial models is necessarily two. That is to say, the exponential quadratic models are indeed the most general class of exponential polynomial models.

In this article, we consider a related class of models, in which the function $H(x, \cdot)$ itself is a polynomial. For instance, in the case where the factor process is scalar-valued, the function

H is of the form

$$H(x,z) = \sum_{k=0}^{n} g_k(x) z^k$$

for $n+1$ differentiable functions $g_k : \mathbb{R}_+ \to \mathbb{R}$. One of the main results is a classification of all such models when H is assumed to satisfy a partial differential equation of the form of equation (1). It turns out that the functions b, a and R are necessarily polynomials of low degree and the functions $(g_k)_k$ solve a system of coupled linear ordinary differential equations. In light of Filipović's maximal degree theorem for exponential polynomial models, it might come as a surprise the degree n is not constrained. We further find necessary and sufficient conditions on the model parameters such that the corresponding stochastic differential equation (3) has a non-explosive solution.

This work is inspired by the interest rate model of Siegel [16]. He showed that for all integers $d \geq 1$ there exist an explicit affine functions R and explicit quadratic functions b, such that the partial differential equation (1) has a solution H such that $H(x,\cdot)$ is affine for all $x \geq 0$. Note that in this case $\partial_{z_i z_j} H$ vanishes identically, and hence the function $a = \sigma\sigma^\top$ need not be specified to verify the partial differential equation. Furthermore, he showed that for a certain choice of σ that the corresponding the stochastic differential equation (3) has a non-explosive solution valued in the bounded state-space

$$I = \{(z_1, \ldots, z_d) : z_i > 0 \text{ for all } i \text{ and } \sum_i z_i < 1\}.$$

We mention also the rational model of Brody & Hughston [1]. Working under the objective measure $\mathbb{P}$, the state price density is modelled $V_t = \alpha(t) + \beta(t) M_t$ where α and β are deterministic functions and M is a $\mathbb{P}$-martingale. We show in section 6 that the Brody–Hughston model is a special case of the time-inhomogeneous polynomial models considered here.

Just as the Brody–Hughston model and the Siegel model described above, most of the polynomial models of this paper (but not all – see section 3.2) have the property that the spot interest rate is *bounded*. This stands in contrast to many familiar models, such as the Vasicek and Cox–Ingersoll–Ross models. Nevertheless, the range of the spot interest rate can be expressed easily in terms of the model parameters, and hence the range can be calibrated to any desired (finite) width.

Finally, a related work is that of Cuchiero, Keller-Ressel & Teichmann [5], who characterise a class of time-homogeneous Markov process Y with the property that the n-th (mixed) moments can be expressed as a polynomial of the initial point Y_0 of degree at most n. Indeed, consider the $d = 1$ case and let F_n be the family of polynomials of degree at most n:

$$F_n = \left\{ f : f(z) = \sum_{k=0}^{n} f_k z^k, \ f_k \in \mathbb{R} \right\}.$$

They study the processes Y that have the property that for *any* degree n and *any* polynomial $g \in F_n$, for all $t \geq 0$ there exists a polynomial $h \in F_n$ such that

$$\mathbb{E}[g(Y_t) | Y_0 = y] = h(y).$$

In contrast, in this work we study processes Z that have the property that for a *fixed* degree n and a *fixed* function R, for all $t \geq 0$ there exists a polynomial $h = H(t, \cdot) \in F_n$ such that

$$\mathbb{E}[e^{-\int_0^t R(Z_s)ds} | Z_0 = z] = h(z).$$

In particular, their results do not imply ours, or vice versa. For further existence results for multi-dimensional polynomial preserving processes, consult the recent paper of Filipović and Larsson [10].

In the remainder of this article is arranged as follows. In section 2, we present one of the main results, a classification of scalar time-homogeneous factor models which satisfy both the analytic assumption that bond prices satisfy a certain partial differential equation and the algebraic assumption that the bond prices can be expressed as a polynomial of the factor. In section 3 we show that a polynomial model in which the spot interest rate is bounded from below necessarily has the property that the spot interest rate is also bounded from above; we also present an explicit example of a polynomial model in which the spot interest rate is unbounded from below. In addition, we provide a complete classification of models satisfying the probabilistic assumption that the corresponding stochastic differential equation has a non-explosive solution valued in a bounded interval. This section contains an easy-to-check formulation of Feller's test of explosion for stochastic differential equations with polynomial coefficients, which might have independent interest. In section 4 we present a spectral representation of the bond prices in the context of a polynomial model. In section 5 we consider a concrete example of this class of models. The Cox–Ingersoll–Ross model is recovered from this example in the limit of degree parameter $n \uparrow \infty$. The properties of this example are analysed and its parameters calibrated to US interest rate data in the case $n = 2$. Finally in section 6, we briefly discuss two extensions: a Hull–White-type extension where the coefficients are allowed to be time dependent, and the higher dimensional case $d > 1$.

2. An algebraic result

This section contains one of the main result of this paper, a classification of models that satisfy the analytic assumption that the pricing function H solves a particular partial differential equation, in addition to having the extra structural property that $H(x, \cdot)$ is a polynomial of fixed degree. To more clearly see the structure of the argument we consider only the time-homogeneous case with $d = 1$ in this section. The time-inhomogeneous and multi-dimensional cases are considered in section 6. The following theorem is of a purely algebraic, rather than probabilistic, nature in the sense the factor process is not mentioned. Related probabilistic results are stated in section 3

Theorem 2.1. *Fix $n \geq 1$, and suppose that*

$$H(x, z) = \sum_{k=0}^{n} g_k(x) z^k \tag{4}$$

for $n + 1$ differentiable functions $g_k : \mathbb{R}_+ \to \mathbb{R}$. Furthermore, assume there exists functions b, σ and R such that

$$\partial_x H = b \partial_z H + \frac{1}{2}\sigma^2 \partial_{zz} H - RH \text{ on } \mathbb{R}_+ \times I \tag{5}$$

where I is a non-trivial interval.

If the functions $(g_k)_k$ are linearly independent, then the following holds true:
Case $n = 1$.
(A) $R(z) = R_0 + R_1 z$ and $b(z) = b_0 + b_1 z + b_2 z^2$ where $R_1 = b_2$.
(B) (g_0, g_1) solve the system of linear ordinary differential equations

$$\begin{aligned}
\dot{g}_0 &= -R_0 g_0 + b_0 g_1 \\
\dot{g}_1 &= -R_1 g_0 + (b_1 - R_0) g_1.
\end{aligned}$$

Case $n \geq 2$.
(A) $R(z) = R_0 + R_1 z + R_2 z^2$, $b(z) = b_0 + b_1 z + b_2 z^2 + b_3 z^3$ and $\sigma^2(z) = a_0 + a_1 z + a_2 z^2 + a_3 z^3 + a_4 z^4$ where the coefficients are such that

$$R_2 = \tfrac{n}{2} b_3 = -\tfrac{n(n-1)}{2} a_4 \text{ and } R_1 = n b_2 + \tfrac{n(n-1)}{2} a_3.$$

(B) $(g_0, \ldots, g_n)$ solves the system of linear ordinary differential equations

$$\begin{aligned}
\dot{g}_k =& g_{k-2} \left((k-2) b_3 + \frac{(k-2)(k-3)}{2} a_4 - R_2 \right) \\
&+ g_{k-1} \left((k-1) b_2 + \frac{(k-1)(k-2)}{2} a_3 - R_1 \right) + g_k \left(k b_1 + \frac{k(k-1)}{2} a_2 - R_0 \right) \\
&+ g_{k+1} \left((k+1) b_0 + \frac{k(k+1)}{2} a_1 \right) + g_{k+2} \frac{(k+2)(k+1)}{2} a_0,
\end{aligned}$$

where we interpret $g_{-2} = g_{-1} = g_{n+1} = g_{n+2} = 0$.

Before proceeding to the proof, note that Theorem 2.1 has an obvious converse: If H has the polynomial form of equation (4) for a certain degree n and that the functions R, b and σ^2 have the form given by condition (A) and the functions $(g_k)_k$ solve the system of ordinary differential equations given by condition (B), then the function H solves the partial differential equation (5). The proof is straightforward.

Proof. Let

$$A_k(z) = k b(z) z^{k-1} + \frac{k(k-1)}{2} \sigma^2(z) z^{k-2} - R(z) z^k.$$

Equation (5) holds if and only if the equation

$$\sum_{k=0}^{n} \dot{g}_k(x) z^k = \sum_{k=0}^{n} g_k(x) A_k(z) \tag{6}$$

holds identically.

For all $m \geq 0$, let

$$F_m = \left\{ f : I \to \mathbb{R} : f(z) = \sum_{k=0}^{m} f_k z^k,\ f_0, \ldots, f_m \in \mathbb{R} \right\}$$

be the set of polynomial functions of degree at most m. Since I is non-trivial, a polynomial $f \in F_m$ uniquely specifies its coefficients $f_0, \ldots, f_m$.

We first show that if equation (6) holds then the functions $A_k \in F_n$ for all k. To see this, use the assumed linear independence of the functions $(g_k)_k$ to pick $n+1$ points $0 \leq x_0 < \ldots < x_n$ such that the $(n+1) \times (n+1)$ matrix $(g_i(x_j))_{i,j}$ is invertible. By evaluating equation (6)

at the points $(x_j)_j$ and solve for the $A_i(z)$, we see that $A_i(z)$ is a linear combination of monomials z^k of degree at most n.
Case $n=1$. Note that

$$\begin{aligned} R(z) &= -A_0(z) \\ b(z) &= A_1(z) + zR(z). \end{aligned}$$

Since A_0 and A_1 are in F_1, i.e. are affine, then R is affine and b is quadratic. Letting $b(z) = b_0 + b_1 z + b_2 z^2$ and $R(z) = R_0 + R_1 z$ the above system equation implies $b_2 = R_1$. Finally, the identity (6) becomes

$$\dot{g}_0 + \dot{g}_1 z = g_0(R_0 + R_1 z) + g_1(b_0 + (b_1 - zR_0)z).$$

Equating coefficients of z yields the necessity and sufficiency of the system of ordinary differential equations.
Case $n \geq 2$. Note that

$$\begin{aligned} R(z) &= -A_0(z) \\ b(z) &= A_1(z) + zR(z) \\ \sigma^2(z) &= A_2(z) - 2zb(z) + z^2 R(z). \end{aligned}$$

Since the functions A_i are polynomials, so are the functions R, b, and σ^2. On the other hand

$$\begin{aligned} A_n(z) &= nb(z)z^{n-1} + \frac{n(n-1)}{2}\sigma^2(z)z^{n-2} - R(z)z^n \\ &= z^{n-2}\left(nb(z)z + \frac{n(n-1)}{2}\sigma^2(z) - R(z)z^2\right) \in F_n \end{aligned}$$

and, since the term in brackets is a polynomial, we have

$$nb(z)z + \frac{n(n-1)}{2}\sigma^2(z) - R(z)z^2 \in F_2 \subseteq F_4. \tag{7}$$

Similarly, since $A_{n-1} \in F_n$ and $A_{n-2} \in F_n$ we have

$$(n-1)b(z)z + \frac{(n-1)(n-2)}{2}\sigma^2(z) - R(z)z^2 \in F_3 \subseteq F_4 \tag{8}$$

$$(n-2)b(z)z + \frac{(n-2)(n-3)}{2}\sigma^2(z) - R(z)z^2 \in F_4. \tag{9}$$

Since

$$\begin{aligned} \sigma^2(z) = &\left(nb(z)z + \frac{n(n-1)}{2}\sigma^2(z) - R(z)z^2\right) + \left((n-2)b(z)z + \frac{(n-2)(n-3)}{2}\sigma^2(z) - R(z)z^2\right) \\ &- 2\left((n-1)b(z)z + \frac{(n-1)(n-2)}{2}\sigma^2(z) - R(z)z^2\right) \end{aligned}$$

inclusions (7), (8) and (9) together yield

$$\sigma^2 \in F_4 \tag{10}$$

Similarly, since

$$zb(z) = \left(nb(z)z + \frac{n(n-1)}{2}\sigma^2(z) - R(z)z^2\right) - \left((n-1)b(z)z + \frac{(n-1)(n-2)}{2}\sigma^2(z) - R(z)z^2\right) - (n-1)\sigma^2$$

inclusions (7), (8) and (10) together yield

$$b \in F_3. \tag{11}$$

Finally, inclusions (7), (10) and (11) together yield

$$R \in F_2.$$

Recall that A_n is of degree at most n. Now substituting $R(z) = \sum_{k=0}^{2} R_k z^k$, $b(z) = \sum_{k=0}^{3} b_k z^k$, $\sigma^2(z) = \sum_{k=0}^{4} a_k z^k$ into the definition of A_n, and setting the coefficient of z^{n+2} to zero yields

$$nb_3 + \frac{n(n-1)}{2}a_4 = R_2 \tag{12}$$

Similarly, equating to zero the coefficient of z^{n+1} in the expansion of A_n yields

$$nb_2 + \frac{n(n-1)}{2}a_3 = R_1.$$

Finally, equating to zero the coefficient of z^{n+1} in the expansion of A_{n-1} yields

$$(n-1)b_3 + \frac{(n-1)(n-2)}{2}a_4 = R_2 \tag{13}$$

Note that equations (12) and (13) together are equivalent to

$$R_2 = \frac{n}{2}b_3 = -\frac{n(n-1)}{2}a_4.$$

Finally, substituting these expressions into equation (6) and comparing the coefficients of the monomials z^k yields the system of ordinary differential equations for the functions for the functions $(g_k)_k$. □

3. Some probabilistic results

In this section we include some results related to the probabilistic assumption that a certain stochastic differential equation has a non-explosive solution.

3.1. **Bounded state space.** In this subsection, we argue that there are good reasons to make the further assumption that the state space I of the factor process Z in a polynomial model is bounded, at least in the one-dimensional case.

Recall that we aim to model the price $P_t(T)$ at time t of a zero-coupon bond of maturity T by the formula $P_t(T) = H(T-t, Z_t)$ where Z_t is the economic factor at time t. Since the payout of the bond is its face value $P_T(T) = 1$, there are economic grounds to assume that the bond prices are bounded. Indeed, to avoid a buy-and-hold arbitrage, one must have $P_t(T) > 0$; furthermore, assuming the existence of a bank account continuously paying the spot interest rate r_t and assuming that $r_t \geq -C$ for some constant $C > 0$, then no arbitrage would imply that $P_t(T) \leq e^{(T-t)C}$.

Continuing with this argument, suppose that the function $H(x, \cdot)$ is bounded on the state space $I \subseteq \mathbb{R}^d$ for each $x \geq 0$. Furthermore, suppose that $H(x, \cdot)$ is a non-constant polynomial. If the dimension $d = 1$, then by an elementary fact of real analysis we can conclude that I is bounded.

We summarise the mathematical content of the above argument in the following proposition:

Proposition 3.1. *Let $I \subseteq \mathbb{R}$ be a non-trivial interval. Suppose that for each $z \in I$, there exists an I-valued weak solution Z of the stochastic differential equation*

$$dZ_t = b(Z_t)dt + \sigma(Z_t)dW_t, \quad Z_0 = z.$$

Furthermore, suppose that there exists a function R such that

$$H(x, z) = \mathbb{E}[e^{-\int_0^x R(Z_s)ds} | Z_0 = z]$$

for all $x \geq 0$. Assume that R has the property that there exists a constant $C > 0$ such that

$$R(z) \geq -C \text{ for all } z \in I.$$

If $H(x, \cdot)$ is a non-constant polynomial for some $x > 0$ then the interval I is bounded.

Remark 3.2. Notice that the assumption that the spot rate $r_t = R(Z_t)$ is bounded *from below* implies that the bond prices are bounded from above. In the case of a one-dimensional polynomial model, boundedness of $H(x, \cdot)$ further implies the boundedness of the state space I. Finally, the boundedness of I and the continuity of R together the that spot interest rate is also bounded *from above.*

On the other hand, notice that the boundedness of $H(x, \cdot)$ does not imply the boundedness of I in the case of *exponential* polynomial models such as Cox–Ingersoll–Ross.

Remark 3.3. Note that in higher dimensions, non-constant polynomials may be bounded on unbounded sets. For instance, consider the polynomial

$$h(z_1, z_2) = z_1 - z_2$$

on the set

$$I = \{(z_1, z_2) : |z_1 - z_2| \leq 1\}.$$

We now mention a pleasant and well-known probabilistic consequence of the boundedness assumption. We state it here in the d dimensional case.

Proposition 3.4. *Let $I \subset \mathbb{R}^d$ be bounded. Suppose the function $H \in C^{1,2}$ satisfies the partial differential equation*

$$\partial_x H = \sum_{1 \leq i \leq d} b_i \, \partial_{z_i} H + \frac{1}{2} \sum_{1 \leq i,j \leq d} a_{ij} \, \partial_{z_i z_j} H - RH \text{ on } \mathbb{R}_+ \times I$$

with boundary condition

$$H(0, z) = 1 \text{ for all } z \in I;$$

and suppose that for each $z \in I$, there exists an I-valued weak solution Z of the stochastic differential equation

$$dZ_t = b(Z_t)dt + \sigma(Z_t)dW_t, \quad Z_0 = z.$$

where $a = \sigma\sigma^\top$. If R is continuous then

$$H(x, z) = \mathbb{E}[e^{-\int_0^x R(Z_s)ds} | Z_0 = z]$$

for all $x \geq 0$ and $z \in I$.

Proof. Fix $z \in I$, and let Z solve the stochastic differential equation with $Z_0 = z$. Also fix a time horizon $x > 0$. As mentioned in the introduction, since H satisfies the partial differential equation, then by Itô's formula we know that the process $M = (M_t)_{0 \leq t \leq x}$ defined by

$$M_t = e^{-\int_0^t R(Z_s)ds} H(x - t, Z_t)$$

is a local martingale. Assuming that the state space I is bounded and that $Z_t \in I$ for all $t \geq 0$, then process M is bounded by a constant by the continuity of R and H.

In particular, the bounded local martingale M is a true martingale by the dominated convergence theorem, and hence

$$\begin{aligned} H(x, z) = M_0 &= \mathbb{E}[M_x] \\ &= \mathbb{E}[e^{-\int_0^x R(Z_s)ds}] \end{aligned}$$

as desired. □

3.2. **An unbounded example.** The message of Proposition 3.1 is that the assumptions that the spot interest rate is bounded from below and that the bond prices are polynomials in a scalar factor together imply that the spot interest rate is bounded from above as well.

We note that many popular interest rate models respect the assumption that the spot rate is bounded from below – an example is the Cox–Ingersoll–Ross model. However, one should mention that some models, such as Vasicek, place no such lower bounds. In particular, while having a lower bound may be a reasonable and desirable feature, but we must grant that this assumption is not universally enforced.

Therefore, we briefly explore the one-dimensional polynomial model, where we drop the assumption that the function R is bounded below on I. We will see via the following example that we can no longer conclude that the state space I is bounded.

Let the state space be $I = (0, \infty)$, and the degree be $n = 2$, the coefficient functions be given by

$$a(z) = z^2, \;\; b(z) = -\frac{1}{2}z^2, \;\; R(z) = -z,$$

and the bond pricing function be

$$H(x, z) = 1 + xz + \frac{1}{2}(e^x - x - 1)z^2.$$

Note that

$$\partial_x H = b\partial_z H + \frac{1}{2}a\partial_{zz}H - RH$$

with initial condition

$$H(0, z) = 1$$

so we are in the setting of Theorem 2.1. By Itô's formula, we know that *if* there exists a solution Z of the stochastic differential equation

$$dZ_t = -\frac{1}{2}Z_t^2 \; dt + Z_t \; dW_t, \;\; Z_0 = z \tag{14}$$

then the process

$$M_t = H(T - t, Z_t)e^{-\int_0^T R(Z_t)dt}$$

is a local martingale.

However, in this case, it is straightforward to show that the unique solution of the stochastic differential equation (14) is given by the formula

$$Z_t = \frac{z e^{W_t - t/2}}{1 + \frac{z}{2}\int_0^t e^{W_s - s/2} ds}.$$

Furthermore, the local martingale M is a true martingale thanks to the identity

$$\begin{aligned} \mathbb{E}(e^{\int_0^x Z_s ds}) &= \mathbb{E}\left[\left(1 + \frac{z}{2}\int_0^x e^{W_s - s/2} ds\right)^2\right] \\ &= H(x, z) \end{aligned}$$

which can be verified by explicit calculation. Note that there is no contradiction with Proposition 3.1 because $\inf_{z \in I} R(z) = -\infty$ in this case.

Remark 3.5. We mention here an interesting (though a bit tangential) observation regarding the above example. It is easy to see that if Z satisfies the stochastic differential equation (14) then the process $Y = e^Z$ defines a local martingale. Indeed, it satisfies the stochastic differential equation

$$dY_t = Y_t \log(Y_t)\ dW_t, \quad Y_0 = e^z.$$

It is slightly less obvious that the process Y is a strictly local martingale. See, for instance, the paper of Goodman [11] for further results.

3.3. **A form of Feller's test.** As argued in Section 3.1, there are economic reasons to consider polynomial models in which the factor process takes values in a bounded interval. Therefore, in this subsection we consider solutions to the scalar stochastic differential equation

$$dZ_t = b(Z_t)dt + \sigma(Z_t)dW_t,$$

which live in a bounded state space $I = (z_{\min}, z_{\max})$. Furthermore, in light of Theorem 2.1, we assume that the coefficients b and σ^2 are polynomials.

To avoid trivial complications, we assume

$$\sigma(z_{\min}) = 0 = \sigma(z_{\max}) \text{ and } \sigma(z) > 0 \text{ for } z_{\min} < z < z_{\max}.$$

Note that the coefficient b is Lipschitz on the closed interval $[z_{\min}, z_{\max}]$, while the coefficient σ is Lipschitz on any interval $[z_{\min} + 1/N, z_{\max} - 1/N]$ for $N > 1$ large enough. Therefore, for every $z \in (z_{\min}, z_{\max})$ the stochastic differential equation has a unique nested family of strong solutions $(Z_{t,N})_{t \in [0, T_N]}$ with $Z_{0,N} = z$, where

$$T_N = \inf\{t > 0 : Z_t \notin (z_{\min} + 1/N, z_{\max} - 1/N)\}.$$

The explosion time T is then defined as

$$T = \sup_N T_N.$$

We are interested in the case where the solution is non-explosive in the sense that $T = \infty$ almost surely.

The classical necessary and sufficient conditions on the functions b and σ is Feller's test of explosion. See for instance Chapter 5 of Karatzas and Shreve's [13] book for an account. Motivated by Theorem 2.1 we adapt Feller's test to the case where the functions b and σ^2

are polynomials. It is likely that the following result is well-known, but we were unable to locate a reference in the literature.

Theorem 3.6. *Let b and a be polynomials. Furthermore, assume*

$$a(z_{\min}) = 0 = a(z_{\max}) \text{ and } a(z) > 0 \text{ for } z_{\min} < z < z_{\max}.$$

Letting $\sigma = \sqrt{a}$, there exists a unique non-explosive strong solution Z of the stochastic differential equation

$$dZ_t = b(Z_t)dt + \sigma(Z_t)dW_t$$

taking values in the interval $(z_{\min}, z_{\max})$ if and only if

$$2b(z_{\min}) - a'(z_{\min}) \geq 0 \geq 2b(z_{\max}) - a'(z_{\max}).$$

Proof. Recall that Feller's test is

$$\mathbb{P}(T = \infty) = 1 \Leftrightarrow v(z_{\min}) = \infty = v(z_{\max}),$$

where Feller's test function is defined by

$$v(x) = \int_{z=c}^{x} \int_{y=z}^{x} \frac{1}{a(z)} e^{\int_y^z \frac{2b(w)}{a(w)} dw} dy\ dz, \text{ for } z_{\min} < x < z_{\max},$$

where $c = \frac{1}{2}(z_{\min} + z_{\max})$. It is enough to consider the behaviour of v near $x = z_{\min}$, as the behaviour near $x = z_{\max}$ is analogous.

By changing variables, we now study the cases where the integral

$$v(z_{\min}) = \int_{z_{\min}}^{c} \frac{p(z)}{a(z)p'(z)} dz$$

is finite or infinite, where

$$p(z) = \int_{z_{\min}}^{z} e^{\int_y^c \frac{2b(u)}{a(u)} du} dy.$$

is the related to the scale function. Now, by assumption the functions a and b are polynomials, and hence near $z_{\min}$ can be written as

$$a(t + z_{\min}) = \alpha t^{A+1} + O(t^{A+2})$$
$$b(t + z_{\min}) = \beta t^{B} + O(t^{B+1})$$

for constants $\alpha > 0$ and $\beta \neq 0$ and for integers $A, B \geq 0$. Note that with this notation

$$2b(z_{\min}) - a'(z_{\min}) = \beta \mathbf{1}_{\{B=0\}} - \alpha \mathbf{1}_{\{A=0\}}.$$

Hence, we must show that $v(z_{\min}) = \infty$ on

$$\{A > 0, B = 0, \beta > 0\} \cup \{A > 0, B > 0\} \cup \{A = 0, B = 0, 2\beta \geq \alpha\}$$

and that $v(0) < \infty$ on the complement

$$\{A > 0, B = 0, \beta < 0\} \cup \{A = 0, B > 0\} \cup \{A = 0, B = 0, 2\beta < \alpha\}.$$

We have the calculation

$$\int_{t+z_{\min}}^{c} \frac{2b(s)}{a(s)} ds = \begin{cases} \text{const} + O(t) & \text{if } B \geq A + 1 \\ -\frac{2\beta}{\alpha} \log t + \text{const} + O(t) & \text{if } B = A \\ \frac{2\beta}{\alpha(A-B)} t^{-(A-B)} + O(t^{1-A+B}) & \text{if } B \leq A - 1 \end{cases}$$

and hence

$$p'(t+z_{\min}) = \begin{cases} \text{const}(1+O(t)) & \text{if } B \geq A+1 \\ t^{-2\beta/\alpha}(\text{const}+O(t)) & \text{if } B = A \\ e^{\frac{2\beta}{\alpha(A-B)}t^{-(A-B)}}(1+O(t)) & \text{if } B \leq A-1 \end{cases}$$

and therefore

$$\frac{p(t+z_{\min})}{a(t+z_{\min})p'(t+z_{\min})} = \begin{cases} \frac{1}{\alpha}t^{-A}(1+O(t)) & \text{if } B \geq A+1 \\ \infty & \text{if } B = A, 2\beta \geq \alpha \\ \frac{1}{2\beta}t^{-A}(1+O(t)) & \text{if } B = A, 2\beta < \alpha \\ \infty & \text{if } B \leq A-1, \beta > 0 \\ \frac{1}{2\beta}t^{-B}(1+O(t)) & \text{if } B \leq A-1, \beta < 0. \end{cases}$$

From this, we see that $v(z_{\min}) = \infty$ precisely on

$$\{B \geq A+1, A \geq 1\} \cup \{B = A, 2\beta \geq \alpha\} \cup \{B = A \geq 1, 2\beta < \alpha\} \\ \cup \{A \geq B+1, \beta > 0\} \cup \{A \geq B+1 \geq 2, \beta < 0\}$$

from which the conclusion follows. □

3.4. **A canonical parametrisation of scalar polynomial models.** We are now in a position to summarise the above results to characterise the range of admissible parameters for which there exists a scalar polynomial model in which the factor process takes values in a bounded open interval I. Since we can replace the state variable Z with $\hat{Z} = \rho(Z)$ and the interest rate function R with $\hat{R} = R \circ \rho^{-1}$, where ρ is an affine transformation, there is no loss of generality in fixing the state space I to be any given interval. Therefore, to simplify some calculations, in this section we will set $I = (-1, 1)$ and will refer to this as the canonical state space in the sequel.

In light of Theorem 2.1, we are interested in necessary and sufficient conditions on the parameters $b_0, \ldots, b_3, a_0, \ldots, a_4$ where the stochastic differential equation

$$dZ_t = b(Z_t)dt + \sigma(Z_t)dW_t$$

has a non-explosive solution valued in the open interval $(-1, 1)$ where

$$b(z) = b_0 + b_1 z + b_2 z^2 + b_3 z^3$$
$$\sigma^2(z) = a_0 + a_1 z + a_2 z^2 + a_3 z^3 + a_4 z^4.$$

In order to enforce the condition $\sigma(-1) = 0 = \sigma(1)$ we rewrite σ^2 as

$$\sigma^2(z) = (1 - z^2)(c_0 + c_1 z + c_2 z^2).$$

In order to describe the parameter space, we first let

$$\mathcal{C} = \{(c_0, c_1, c_2) : c_0 + c_1 z + c_2 z^2 > 0 \text{ for all } -1 < z < 1\}. \tag{15}$$

This set can be described more explicitly:

Proposition 3.7.

$$\mathcal{C} = \{c_0 > 0, -c_0 \leq c_2 \leq c_0, |c_1| \leq c_0 + c_2\} \cup \{c_0 > 0, c_2 > c_0, |c_1| < 2\sqrt{c_0 c_2}\}.$$

Proof. For a fixed triplet (c_0, c_1, c_2) let $c(z) = c_0 + c_1 z + c_2 z^2$. Suppose that $c(z) > 0$ for all $-1 < z < 1$. Note that $c(0) = c_0$, so $c_0 > 0$ necessarily.

By continuity, we have $c(-1) = c_0 - c_1 + c_2 \geq 0$ and $c(1) = c_0 + c_1 + c_2 \geq 0$. Hence $c_2 \geq -c_0$ and $|c_1| \leq c_0 + c_2$ necessarily. Note that

$$\begin{aligned} f(z) &\geq c_0 - |c_1||z| + c_2 z^2 \\ &\geq c_0 - (c_0 + c_2)|z| + c_2 z^2 \\ &= (c_0 - |z|c_2)(1 - |z|). \end{aligned}$$

From this we see that if $c_2 \leq c_0$ then $c(z) > 0$ whenever $|z| < 1$, and hence the condition $|c_1| \leq c_0 + c_2$ is also sufficient.

So now suppose that $c_2 > c_0$. Letting

$$z_0 = \sqrt{\frac{c_0}{c_2}}$$

we have $0 < z_0 < 1$ and

$$c(z_0) = (c_1 + 2\sqrt{c_0 c_2}) z_0.$$

Hence the condition $c_1 > -2\sqrt{c_0 c_2}$ is necessary. By considering $c(-z_0)$ we see that the condition $c_1 < 2\sqrt{c_0 c_2}$ is also necessary. Finally, writing

$$c(z) = c_0 - \frac{c_1^2}{4c_2} + c_2 \left(z - \frac{c_1}{2c_2} \right)^2.$$

we see that the condition $|c_1| < 2\sqrt{c_0 c_2}$ is sufficient as well. $\square$

We now consider the function b. Let

$$\mathcal{B}_{c_0,c_1,c_2} = \{(b_0, b_1, b_2, b_3) : 2b(-1) - a'(-1) \geq 0 \geq 2b(1) - a'(1)\}. \tag{16}$$

Again, this set can be described more explicitly:

Proposition 3.8.

$$\mathcal{B}_{c_0,c_1,c_2} = \{|b_0 + b_2 + c_1| \leq -(b_1 + b_3 + c_0 + c_2)\}$$

The proof is straightforward.

We can now parametrise the general scalar, time-homogeneous polynomial model with factor process taking values in $(-1, 1)$. For each $n \geq 2$, we let

$$\begin{aligned} \mathcal{P}_n = \{&(R_0, R_1, R_2, b_0, b_1, b_2, b_3, c_0, c_1, c_2) : (c_0, c_1, c_2) \in \mathcal{C}, (b_0, b_1, b_2, b_3) \in \mathcal{B}_{c_0,c_1,c_2}, \\ &R_1 = nb_2 - \tfrac{1}{2}n(n-1)c_1, R_2 = (n-1)b_3 = \tfrac{1}{2}(n-1)(n-2)c_2 \big\} \end{aligned} \tag{17}$$

To summarise, we have proven that if

$$(R_0, R_1, R_2, b_0, b_1, b_2, b_3, c_0, c_1, c_2) \in \mathcal{P}_n$$

then for every $z \in (-1, 1)$ the stochastic differential equation

$$dZ_t = (b_0 + b_1 Z_t + b_2 Z_t^2 + b_3 Z_t^3)dt + \sqrt{(1 - z^2)(c_0 + c_1 Z_t + c_2 Z_t^2)} dW_t, Z_0 = z$$

has a unique strong non-explosive solution taking values in $(-1,1)$. Furthermore, the solution has the property that

$$\mathbb{E}[e^{-\int_0^x (R_0+R_1Z_s+R_2Z_s^2)ds}|Z_0=z]=\sum_{k=0}^{n} g_k(x)z^k$$

where the functions $g_0,\dots,g_k$ solve the the system of linear ordinary differential equations

$$\begin{aligned}\dot{g}_k =& g_{k-2}\left((k-2)b_3+\frac{(k-2)(k-3)}{2}a_4-R_2\right)\\ &+g_{k-1}\left((k-1)b_2+\frac{(k-1)(k-2)}{2}a_3-R_1\right)+g_k\left(kb_1+\frac{k(k-1)}{2}a_2-R_0\right)\\ &+g_{k+1}\left((k+1)b_0+\frac{k(k+1)}{2}a_1\right)+g_{k+2}\frac{(k+2)(k+1)}{2}a_0,\end{aligned}$$

where we interpret $g_{-2}=g_{-1}=g_{n+1}=g_{n+2}=0$, subject to the boundary condition

$$g_0(0)=1, g_1(0)=\cdots=g_n(0)=0.$$

4. A spectral representation

We are in the setting of the scalar time-homogeneous polynomial model with the factor process taking values in a bounded open interval I. The coefficient functions $(g_k)_k$ are the solution of the system of differential equations which can be equivalently described as follows. Let $S=(S_{i,j})_{i,j=0}^n$ be the $(n+1)\times(n+1)$ matrix with entries

$$S_{j+k,j}=jb_{k+1}+\tfrac{j(j-1)}{2}a_{k+2}-R_k$$

and where $R_k=b_k=a_k=0$ when $k<0$ and $R_k=b_{k+1}=a_{k+2}=0$ when $k>2$. For instance, when $n\geq 4$, the matrix has the form

$$S=\begin{pmatrix} -R_0 & b_0 & a_0 & & & \\ -R_1 & b_1-R_0 & 2b_0+a_1 & 3a_0 & & \\ -R_2 & b_2-R_1 & 2b_1+a_2-R_0 & 3b_0+3a_1 & 6a_0 & \\ & b_3-R_2 & 2b_2+a_3-R_1 & 3b_1+3a_2-R_0 & 4b_0+6a_1 & \ddots \\ & & 2b_3+a_4-R_2 & 3b_2+3a_3-R_1 & 4b_1+6a_2-R_0 & \ddots \\ & & & \ddots & \ddots & \ddots \end{pmatrix}.$$

Now letting

$$G(x)=(g_0(x),g_1(x),\dots,g_n(x))^\top$$

The system of differential equations becomes

$$\dot{G}=SG,$$

and, in particular, the solution can be expressed as

$$G(x)=e^{Sx}G(0),$$

where the boundary condition for zero-coupon bond pricing is given by

$$G(0)=(1,0,\dots,0)^\top.$$

It turns out that the matrix S has a nice property:

Proposition 4.1. *The eigenvalues $\lambda_0, \ldots, \lambda_n$ of S are real and satisfy*

$$\lambda_i \leq - \inf_{z \in I} R(z).$$

for all i.

To prove Proposition 4.1, we first prove a result on the existence of an invariant measure which may have independent interest.

Proposition 4.2. *Suppose (c_0, c_1, c_2) is in the interior of the set $\mathcal{C}$ defined by equation (15) and that b is in the set $\mathcal{B}_{c_0,c_1,c_2}$ defined by equation (16) . Then there exists a positive, integrable function f satisfying the differential equation*

$$bf = \frac{1}{2}(af)'$$

with boundary conditions

$$\lim_{z \downarrow -1} a(z)f(z) = 0 = \lim_{z \uparrow 1} a(z)f(z),$$

where

$$b(z) = b_0 + b_1 z + b_2 z^2 + b_3 z^3$$
$$a(z) = (1 - z^2)(c_0 + c_1 z + c_2 z^2).$$

Remark 4.3. If the function f is normalised so that $\int_{-1}^{1} f(z)dz = 1$, then f is the unique invariant density for the diffusion Z with drift b and volatility σ. That is, if the initial condition Z_0 is distributed with density f, then Z_t has the same distribution for all $t \geq 0$.

Proof. Since (c_0, c_1, c_2) is in the interior of $\mathcal{C}$, we may assume that either

$$-c_0 \leq c_2 \leq c_0 \text{ and } |c_1| < c_0 + c_2$$

or

$$c_2 > c_0, |c_1| < 2\sqrt{c_0 c_2}\}.$$

In either case, the function $c(z) = c_0 + c_1 z + c_2 z^2$ is strictly positive on the closed interval $[-1, 1]$. In particular, the function $a(z) = (1 - z^2)c(z)$ has simple roots at $z = -1$ and $z = 1$ with

$$a'(-1) = 2c(-1) > 0 > a'(1) = -2c(1).$$

Furthermore, since $(b_0, b_1, b_2, b_3) \in \mathcal{B}_{(c_0,c_1,c_2)}$, we have

$$b(-1) > 0 > b(1).$$

Now any positive solution to the differential equation is of form

$$f(z) = \frac{C}{a(z)} e^{\int_0^z \frac{2b(s)}{a(s)} ds}.$$

for $|z| < 1$, where $C > 0$ is a constant.

Following the proof of Theorem 3.6 we focus on the left-hand end point $z = -1$. We must show that such an f is integrable and $tf(t - 1) \to 0$ as $t \downarrow 0$. Writing

$$b(t - 1) = \beta + O(t)$$
$$a(t - 1) = \alpha t + O(t^2)$$

a routine calculation shows that

$$f(t-1) = \frac{C}{\alpha} t^{\frac{2\beta}{\alpha}-1}$$

from which the conclusion follows. $\square$

We can now prove Proposition 4.1.

Proof. We can and do assume that have the made the canonical choice of state space $I = (-1, 1)$ introduced in section 3.4. Since S varies continuously with the model parameters, there is no loss of generality to assume that the parameters are in the interior of the $\mathcal{P}_n$ defined by equation (17). By Proposition 4.2 there exists an invariant density f.

Consider the inner product on $\mathbb{R}^{n+1}$ defined by

$$\begin{aligned} \langle p, q \rangle &= \sum_{i=0}^{n} \sum_{j=0}^{n} p_i q_j \int_{-1}^{1} z^{i+j} f(z) dz \\ &= \int_{-1}^{1} \hat{p}(z) \hat{q}(z) f(z) dz \end{aligned}$$

where $\hat{}$ is the linear operator sending the vector

$$p = (p_0, \ldots, p_n)$$

to the $n+1$ degree polynomial

$$\hat{p}(z) = \sum_{k=0}^{n} p_k z^k.$$

The key observation is that

$$\widehat{Sp} = \frac{1}{2} a \hat{p}'' + b \hat{p}' - R \hat{p}.$$

Note that

$$\begin{aligned} \langle p, Sq \rangle &= -\int_{-1}^{1} [\tfrac{1}{2} a \hat{p}' \hat{q}' + R \hat{p} \hat{q}] f \ dz \\ &= \langle Sp, q \rangle \end{aligned}$$

by integration by parts, where we have used the boundary condition

$$\lim_{z \downarrow -1} a(z) f(z) = 0 = \lim_{z \uparrow 1} a(z) f(z).$$

In particular, we see that S is symmetric with respect to this inner product and hence all eigenvalues are real. The inequality

$$\begin{aligned} \langle p, Sp \rangle &\leq -\int_{-1}^{1} R \hat{p}^2 f \ dz \\ &\leq -\inf_{-1<z<1} R(z) \langle p, p \rangle \end{aligned}$$

implies the claimed upper bound on the spectrum. $\square$

Remark 4.4. Of course, the eigenvalues of the matrix S are the zeros of the characteristic polynomial which has degree $n+1$. It is well know that there exists an explicit formula, discovered by Ferrari in 1540, for the zeros of quartic polynomials, and hence the eigenvalues of S can be expressed in a closed formula in terms of the matrix entries when $n \leq 3$. In particular, in this case, the bond pricing function H can be written, at least in principle, explicitly in terms of the model parameters.

When $n \geq 4$, there is little hope for explicit formulae for the function H in terms of the model parameters. However, note that the matrix is sparse, in the sense that there are at most five non-zero matrix entries per row. In particular, the product of the matrix exponential e^{Sx} and the vector $(1, 0, \ldots, 0)^\top$ can be computed efficiently, and hence the lack of explicit formulae is not necessarily a prohibitive disadvantage.

The proof of Proposition 4.1 shows that when there exists an invariant density f, then

$$S^\top M = MS$$

where $M = (M_{ij})_{ij}$ is the $(n+1) \times (n+1)$ positive definite matrix with entries

$$M_{ij} = \int_I z^{i+j} f(z) dz.$$

Suppose that the $n+1$ real eigenvalues of the matrix S are $\lambda_0, \ldots, \lambda_n$. Then the matrix S has the spectral decomposition

$$S = \sum_{i=0}^{n} \lambda_i \; u_i \; v_i$$

where u_i is the right-eigenvector and v_i the left-eigenvector associated to the eigenvalue λ_i, scaled such that

$$v_i u_j = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j. \end{cases}$$

For convenience, we choose the normalisation

$$u_i^\top M u_i = 1,$$

and note that the left- and right-eigenvectors are related by

$$v_i = u_i^\top M.$$

Now given the ith right-eigenvector u_i we can form the polynomial $\hat{u}_i(z) = \sum_{k=0}^{n} u_{i,k} z^k$. Note that $\hat{u}_i$ is an eigenfunction of a certain differential operator

$$\left(b(z)\partial_z + \tfrac{1}{2} a(z) \partial_{zz} - R(z) \right) \hat{u}_i(z) = \lambda_i \hat{u}_i(z),$$

and that the ith left-eigenvector v_i is related to $\hat{u}_i$ by the formula

$$v_{i,k} = \int_I z^k \hat{u}_i(z) f(z) dz$$

In particular, the bond pricing function takes the form

$$H(x, z) = \sum_{i=0}^{n} Q_i(z) e^{\lambda_i x}$$

where the function Q_i is the (at most) n degree polynomial

$$Q_i(z) = \hat{u}_i(z) \int_I \hat{u}_i(s) f(s) ds$$

That is to say, the bond price can be seen to be a linear combination of the bond prices arising from $n+1$ models with constant interest rates $r = -\lambda_i$, where the coefficients Q_i of the combination depend on the factor process. Note that by setting $x = 0$ we have

$$\sum_{i=0}^{n} Q_i(z) = 1$$

so it is tempting to think of the numbers $(Q_i(z))_i$ as probabilities; however, in general $Q_i(z) < 0$ for some i and $z \in I$, so such an interpretation is not always valid. An example where Q_i takes negative values can be found in section 5 below.

In the general case, where the parameters are such that no invariant density exists, the matrix S is not necessarily diagonalisable. In this case, the bond pricing formula must be modified to

$$H(x, z) = \sum_{i=0}^{n} Q_i(x, z) e^{\lambda_i x} \tag{18}$$

where now the weight functions Q_i are polynomials in both x and z and can be computed from the Jordan decomposition. An example where the matrix S is not diagonalisable is discussed in Section 3.2 – though strictly speaking, the setting is slightly different there since the state space of the stochastic differential equation (14) is unbounded.

One consequence of formula (18) is that the long maturity interest rate can calculated as

$$\lim_{x \to \infty} -\frac{1}{x} \log H(x, z) = -\max_i \lambda_i.$$

for all $z \in I$, unless the coefficient $Q_i(x, z)$ of the maximum eigenvalue is identically zero.

5. An example

In this section we explore a concrete realisation of a polynomial model. The purpose of this account is as a proof of concept and is not intended as an endorsement of this particular model over others.

In the general polynomial framework, the function $R : I \to \mathbb{R}$, mapping the factor process to the spot interest rate, is a quadratic function. In the following example, we assume that R is affine. By an affine change of variables, we can and will take the spot rate itself as the factor process. Note that this choice of parametrisation differs from the canonical choice introduced in Section 3.4.

We consider here a model where the spot rate is the solution of the stochastic differential equation

$$dr_t = \alpha(\beta - r_t)dt + \sqrt{r_t \left(\gamma - \tfrac{2}{n} r_t\right)\left(\gamma - \tfrac{1}{n-1} r_t\right)} dW_t.$$

This model is inspired by the Cox–Ingersoll–Ross: the parameter $\beta > 0$ plays the role of a long time mean level, the parameter $\alpha > 0$ controls the rate of mean-reversion, and the parameter $\gamma > 0$ is related to the spot rate volatility, in the sense that the infinitesimal

variance of dr_t is approximately $\gamma^2 r_t dt$ when r_t is very small. Indeed, sending $n \uparrow \infty$ formally recovers the Cox–Ingersoll–Ross stochastic differential equation.

However, unlike the Cox–Ingersoll–Ross model, the interest rate stays within the bounded interval $I = (0, \frac{1}{2}\gamma n)$. Applying Theorem 3.6, we see that the appropriate Feller conditions are

$$2\alpha\beta \geq \gamma^2 \text{ and } \alpha(\gamma n - 2\beta) \geq \gamma^2 \frac{n-2}{n-1}.$$

To further simplify this discussion, we set the degree parameter $n = 2$, so that the set of feasible parameters becomes

$$0 \leq \beta \leq \gamma \leq \sqrt{2\alpha\beta}.$$

The corresponding matrix S introduced in Section 4 takes the form

$$S = \begin{pmatrix} 0 & \alpha\beta & 0 \\ -1 & -\alpha & 2\alpha\beta + \gamma^2 \\ 0 & -1 & -2(\alpha+\gamma) \end{pmatrix}$$

from which bond pricing function can be evaluated efficiently by the formula

$$H(x, r) = (1, r, r^2) e^{Sx} (1, 0, 0)^\top.$$

Notice if the parameters are such that

$$0 < \beta < \gamma \leq \sqrt{2\alpha\beta}$$

then the process is ergodic in the pricing measure $\mathbb{Q}$, and its invariant density is given by the stationary solution of the corresponding Fokker–Planck partial differential equation:

$$\begin{aligned} f(r) &\propto \frac{1}{\sigma(r)^2} e^{\int_0^r \frac{2b(\rho)}{\sigma(\rho)^2} d\rho} \\ &\propto r^{\zeta-1} (\gamma - r)^{-\zeta-2} e^{-\theta/(\gamma - r)}. \end{aligned}$$

where

$$\zeta = \frac{2\alpha\beta}{\gamma^2} \text{ and } \theta = 2\alpha \left(1 - \frac{\beta}{\gamma}\right).$$

To confirm that this model can produce sensible looking yield curves, the parameters α, β, γ have been calibrated to US Treasury constant maturity rates freely available from the Federal Reserve Economic Data (FRED) website [8]. The 251 daily yield curve observations for maturities

$$x_i \in \{1/12, 3/12, 6/12, 1, 2, 3, 5, 7, 10, 20, 30 \text{ years } \}$$

and dates

$$t_j \in \{ \text{ 19 August 2015}, \ldots, \text{ 18 August 2016 } \}$$

were fit to this model. The parameter values

$$\alpha = 0.248, \beta = 3.1\%, \gamma = 12.9\%$$

seem to be a (local) minimum of the penalty function

$$\sum_{i,j} (Y_{\text{obs}}(x_i, t_j) - Y_{\alpha,\beta,\gamma}(x_i, r_j))^2$$

subject to Feller feasibility condition, where the model yield is $Y(x,r) = -\frac{1}{x}\log H(x,r)$ and the spot rate r_j on day t_j is linearly interpolated from the two shortest dated yields via the formula

$$r_j = \frac{3}{2}Y_{\text{obs}}(1/12, t_j) - \frac{1}{2}Y_{\text{obs}}(3/12, t_j).$$

Figure 1 shows the fitted yield curve compared to the observed yield curve on the first and last days of the sample period. It is no surprise that the fit is better for left-hand end, as there are more observations at short maturities. Although the fit is not extremely impressive, it is probably sufficiently good for the modest goal of illustrating the potential utility of the class of polynomial models.

Recall that the bond pricing function has two representations

$$\begin{aligned} H(x,r) &= g_0(x) + g_1(x)r + g_2(x)r^2 \\ &= Q_0(r)e^{\lambda_0 x} + Q_1(r)e^{\lambda_1 x} + Q_2(r)e^{\lambda_2 x} \end{aligned}$$

where the exponential polynomial functions $(g_k)_k$ solve the system

$$\begin{aligned} \dot{g}_0 &= \alpha\beta g_1 \\ \dot{g}_1 &= -g_0 - \alpha g_1 + (2\alpha\beta + \gamma^2)g_2 \\ \dot{g}_2 &= -g_1 - 2(\alpha+\gamma)g_2 \end{aligned}$$

with initial condition

$$g_0(0) = 1, \quad g_1(0) = g_2(0) = 0.$$

and the quadratic functions $(Q_i)_i$ satisfy the eigenequation

$$\frac{1}{2}r(\gamma - r)^2 Q_i''(r) + \alpha(\beta - r)Q_i'(r) - rQ_i(r) = \lambda_i Q_i(r)$$

with the normalisation such that

$$\int_0^\gamma Q_i(r)f(r)dr = \int_0^\gamma Q_i(r)^2 f(r)dr,$$

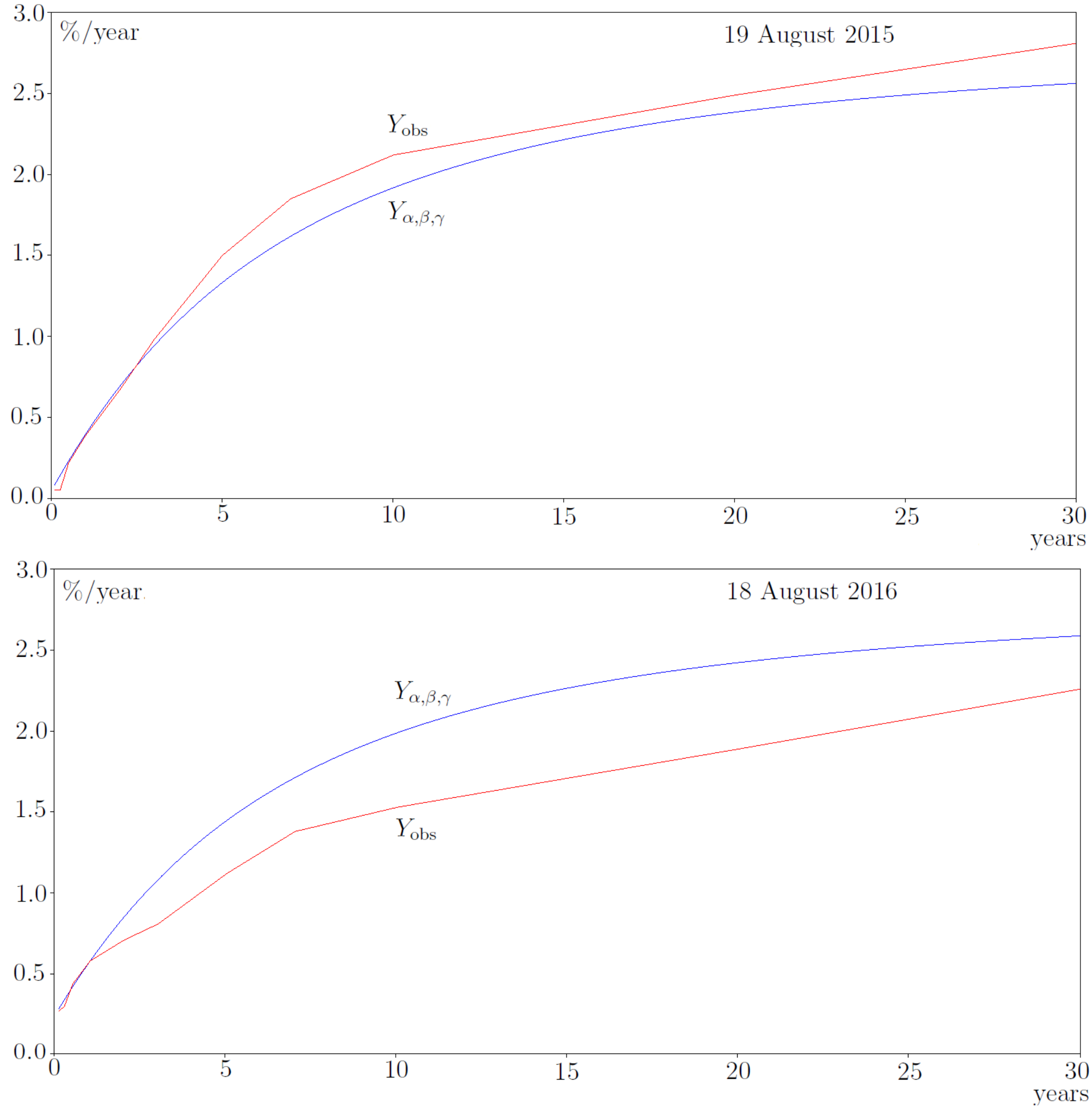
where $(\lambda_i)_i$ are the eigenvalues of the matrix S. Note that $Q_1(r) < 0$ for some values of r. The graphs of the coefficient functions $(g_k)_k$ and $(Q_i)_i$ are shown in Figures 2 and 3.

See the recent thesis [3] of Cheng for other calibrated examples.

6. Extensions

In this section, we will extend theorem 2.1 in two different ways: namely allowing time dependency and allowing a multi-dimensional factor process. We focus on the algebraic result relating the analytic assumption that the bond pricing function satisfies a certain partial differential equation to the assumption that the bond prices are polynomial in the factor process. We omit a discussion of the probabilistic assumption of whether a certain stochastic differential equation has a non-explosive solution.

FIGURE 1.



6.1. **Hull–White-type extension.** As usual, by incorporating time-dependent parameters, we can hope to have a better model calibration. We introduce time dependency both in the dynamics of the factor process $(Z_t)_{t\geq 0}$ and the coefficient functions $(g_k)_k$. As one may expect, we will establish a similar algebraic result in this case.

To be clear, we now consider a factor process $(Z_t)_{t\geq 0}$ to be a non-explosive solution valued in a non-trivial interval $I \subseteq \mathbb{R}$ to the time-inhomogeneous stochastic differential equation

$$dZ_t = b(t, Z_t)dt + \sigma(t, Z_t)dW_t.$$

FIGURE 2.

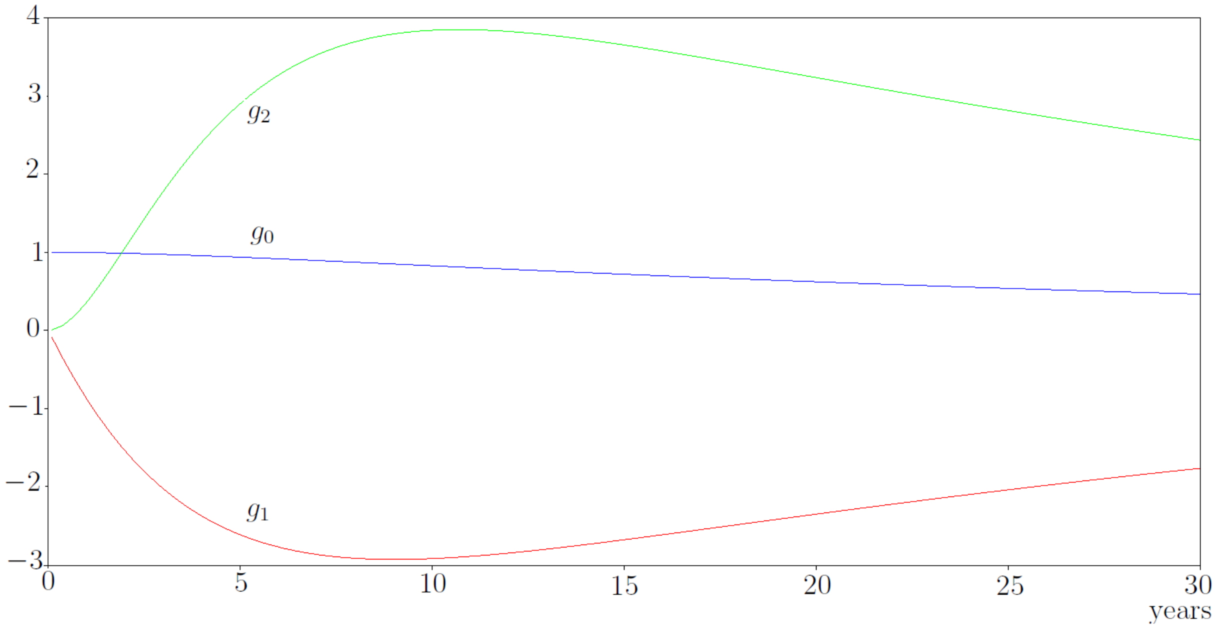
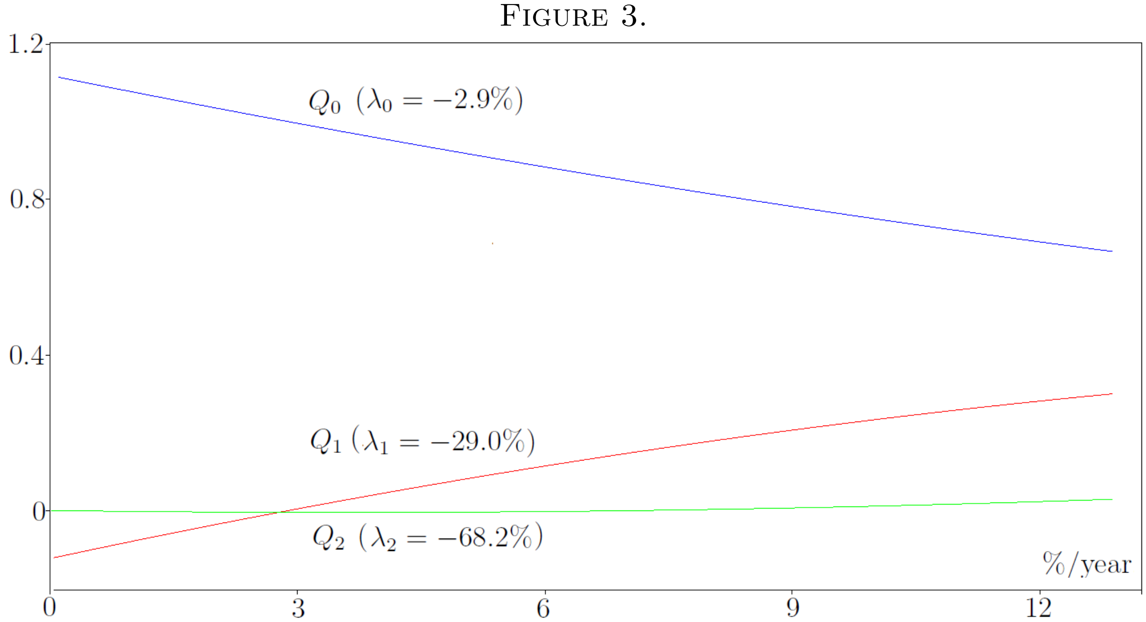


FIGURE 3.



The spot rate is modelled as $r_t = R(t, Z_t)$ and the bond prices are defined by

$$P_t(T) = \sum_{k=0}^{n} g_k(t, T) Z_t^k$$

where $g_k : \Delta \to \mathbb{R}$ are smooth deterministic functions satisfying the boundary conditions:

$$\begin{aligned} g_0(T,T) &= 1 \\ g_k(T,T) &= 0 \quad \text{for all} \quad 1 \leq k \leq n \end{aligned}$$

where $\Delta = \{(t,T) : 0 \leq t \leq T\}$.

By adding the t component, the analytic assumption is this case is that the functions g_k satisfy the partial differential equation

$$\sum_{k=0}^{n} \partial_t g_k(t,T) z^k = \sum_{k=0}^{n} g_k(t,T) A_k(t,z) \ \text{ for all } 0 \leq t \leq T, z \in I \tag{19}$$

where

$$A_k(t,z) = R(t,z)z^k - kb(t,z)z^{k-1} - \frac{k(k-1)}{2}\sigma^2(t,z)z^{k-2}$$

The algebraic result in this case becomes the following:

Theorem 6.1. *Suppose that for each $t \geq 0$, the functions $(g_k(t,\cdot))_k$ are linearly independent and satisfy equation* (19)*.*
Case $n = 1$.
(A) $R(t,z) = R_0(t) + R_1(t)z$ and $b(t,z) = b_0(t) + b_1(t)z + b_2(t)z^2$ where $R_1(t) = b_2(t)$.
(B) (g_0, g_1) solves the system of linear ordinary differential equations

$$\begin{aligned} -\partial_t g_0 &= -R_0 g_0 + b_0 g_1 \\ -\partial_t g_1 &= -R_1 g_0 + (b_1 - R_0) g_1. \end{aligned}$$

Case $n \geq 2$.
(A) $R(t,z) = R_0(t) + R_1(t)z + R_2(t)z^2$, $b(t,z) = b_0(t) + b_1(t)z + b_2(t)z^2 + b_3(t)z^3$ and $\sigma^2(z) = a_0(t) + a_1(t)z + a_2(t)z^2 + a_3(t)z^3 + a_4(t)z^4$ where the coefficients are such that

$$R_2(t) = \tfrac{n}{2}b_3(t) = -\tfrac{n(n-1)}{2}a_4(t) \ \textit{and} \ R_1(t) = nb_2(t) + \tfrac{n(n-1)}{2}a_3(t).$$

(B) $(g_0, \ldots, g_n)$ solves the system of linear ordinary differential equations

$$\begin{aligned} -\partial_t g_k =& g_{k-2}\left((k-2)b_3 + \frac{(k-2)(k-3)}{2}a_4 - R_2\right) \\ &+ g_{k-1}\left((k-1)b_2 + \frac{(k-1)(k-2)}{2}a_3 - R_1\right) + g_k\left(kb_1 + \frac{k(k-1)}{2}a_2 - R_0\right) \\ &+ g_{k+1}\left((k+1)b_0 + \frac{k(k+1)}{2}a_1\right) + g_{k+2}\frac{(k+2)(k+1)}{2}a_0, \end{aligned}$$

where we interpret $g_{-2} = g_{-1} = g_{n+1} = g_{n+2} = 0$.

The proof is essentially the same as that of Theorem 2.1 so is omitted.

6.2. **Brody–Hughston rational model.** In the paper [1] of Brody & Hughston, the following rational model is discussed. Let M be a positive martingale under the *objective measure* $\mathbb{P}$, and suppose $M_0 = 1$. Set

$$V_t = \alpha(t) + \beta(t) M_t$$

where α and β are positive, continuously differentiable, deterministic functions. The idea is that V is a model for the state price density. Therefore, bond prices are given by the formula

$$\begin{aligned} P_t(T) &= \frac{1}{V_t} \mathbb{E}^{\mathbb{P}}(V_T | \mathcal{F}_t) \\ &= \frac{\alpha(T) + \beta(T) M_t}{\alpha(t) + \beta(t) M_t}. \end{aligned}$$

Note that the bond prices are a rational function of the random variable M_t, giving the model its name. Furthermore, by setting $\alpha(t) + \beta(t) = P_0(t)$ for $t \geq 0$, this model can match the initial term structure of interest rates.

On the other hand, notice that we can write the bond prices as

$$P_t(T) = g_0(t,T) + g_1(t,T) Z_t$$

where the coefficients are defined by

$$g_0(t,T) = \frac{\beta(T)}{\beta(t)} \text{ and } g_1(t,T) = \frac{\alpha(T)\beta(t) - \beta(T)\alpha(t))}{\beta(t)}$$

and where we let

$$Z_t = \frac{1}{V_t}$$

be the factor process. In particular, this is an affine factor model and hence should be described by Theorem 6.1. We now carry out the verification under the assumption that

$$dM_t = \nu(t, M_t) M_t dB_t$$

where B is a $\mathbb{P}$-Brownian motion and ν is bounded.

In this framework, we can define the spot rate as

$$\begin{aligned} r_t &= -\partial_T P_t(T)|_{T=t} \\ &= -\frac{\dot{\alpha}(t) + \dot{\beta}(t) M_t}{\alpha(t) + \beta(t) M_t} \\ &= R_0(t) + R_1(t) Z_t \end{aligned}$$

where

$$R_0(t) = -\frac{\dot{\beta}(t)}{\beta(t)} \text{ and } R_1(t) = \frac{\dot{\beta}(t)\alpha(t) - \dot{\alpha}(t)\beta(t)}{\beta(t)}.$$

Note that

$$\begin{aligned} dV_t &= \left(\dot{\alpha}(t) + \dot{\beta}(t) M_t \right) dt + \beta(t) dM_t \\ &= -V_t (r_t dt + \lambda_t dB_t) \end{aligned}$$

where λ_t is the market price of risk defined by

$$\begin{aligned}\lambda_t &= -\frac{\beta(t)\nu(t, M_t)M_t}{\alpha(t) + \beta(t)M_t} \\ &= \nu(t, M_t)(\alpha(t)Z_t - 1)\end{aligned}$$

Since the process $(\lambda_t)_{0\le t\le T}$ is bounded, we can define the equivalent risk-neutral pricing measure $\mathbb{Q}$ by

$$\begin{aligned}\frac{d\mathbb{Q}}{d\mathbb{P}} &= e^{\int_t^T r_s ds} V_t \\ &= e^{-\frac{1}{2}\int_0^T \lambda_t dt + \int_0^T \lambda_t dB_t}\end{aligned}$$

to recover the usual pricing formula

$$P_t(T) = \mathbb{E}^{\mathbb{Q}}[e^{-\int_t^T r_s ds} | \mathcal{F}_t].$$

Finally, we consider the dynamics of the factor process $Z = V^{-1}$. By Itô's formula we have

$$\begin{aligned}dZ_t &= Z_t[(r_t + \lambda_t^2)dt + \lambda_t dB_t] \\ &= (b_1(t)Z_t + b_2(t)Z_t^2)dt + \sigma(t, Z_t)dW_t\end{aligned}$$

where

$$\begin{aligned}b_1(t) &= -\frac{\dot{\beta}(t)}{\beta(t)} \\ b_2(t) &= \frac{\dot{\beta}(t)\alpha(t) - \dot{\alpha}(t)\beta(t)}{\beta(t)} \\ \sigma(t, z) &= \nu\left(t, \frac{1 - z\alpha(t)}{z\beta(t)}\right)(\alpha(t)z - 1)z\end{aligned}$$

and where the process $(W_t)_{0\le t\le T}$ defined by

$$W_t = B_t + \int_0^t \lambda_s ds$$

is a $\mathbb{Q}$ Brownian motion by Girsanov's theorem. In particular, notice that the drift is quadratic in Z and $b_2(t) = R_1(t)$ as predicted by Theorem 6.1, while the volatility is determined by the function ν.

6.3. **Multi-dimensional factor process.** In this subsection, we will extend the polynomial model framework by allowing both the factor process $(Z_t)_{t\ge 0}$ and the background Brownian motion $(W_t)_{t\ge 0}$ to be multi-dimensional. To be more specific, let $(W_t)_{t\ge 0}$ be a D-dimensional Brownian motion. Let $(Z_t)_{t\ge 0}$ be the factor process taking values in $I \subseteq \mathbb{R}^d$, assuming to be the (non-explosive) solution of the stochastic differential equation

$$dZ_t = b(Z_t)dt + \sigma(Z_t)dW_t$$

for some deterministic functions $b : I \to \mathbb{R}^d$ and $\sigma : I \to \mathbb{R}^{d\times D}$. We define the diffusion function $a = \sigma\sigma^\top$, and note that the only role played by the parameter D is as the upper bound on the rank of the matrix $a(z)$.

For $k = (k_1, \ldots, k_d) \in \mathbb{Z}_+^d$ and $z = (z_1, \ldots, z_d) \in \mathbb{R}^d$, we define the monomial z^k as follows:

$$z^k = z_1^{k_1} \cdots z_d^{k_d}$$

For each $m \geq 0$, we define the following set of indices

$$K_m = \{k \in \mathbb{Z}_+^d : k_1 + \ldots + k_d \leq m\}.$$

With this notation, we consider models where the bond price is of the form

$$P_t(T) = \sum_{k \in K_n} g_k(T - t) Z_t^k$$

where the functions g_k satisfy the boundary conditions

$$g_k(0) = 1 \text{ if } k = (0, \ldots, 0)$$
$$g_k(0) = 0 \text{ otherwise}$$

The spot rate is modelled as $r_t = R(Z_t)$ for some deterministic function $R : I \to \mathbb{R}$.

The partial differential equation (1) gives rise to the condition

$$\sum_{k \in K_n} \dot{g}_k(x) z^k = \sum_{k \in K_n} g_k(x) A_k(z) \tag{20}$$

holds for any $x \geq 0$ and $z \in I$, where the functions A_k are defined as

$$A_k(z) = \sum_{1 \leq i \leq d} b_i(z) \partial_{z_i} z^k + \frac{1}{2} \sum_{1 \leq i,j \leq d} a_{ij}(z) \partial_{z_i z_j} z^k - R(z) z^k.$$

Finally we define the notation

$$F_m = \left\{ f : I \to \mathbb{R}, f(z) = \sum_{k \in K_m} f_k z^k, \quad f_k \in \mathbb{R} \right\}$$

to be the family of polynomials in d variables of total degree less or equal to m. We will assume that the state space $I \subseteq \mathbb{R}^d$ is non-trivial in the sense that $f \in F_m$ uniquely determines the coefficients $(f_k)_k$, or more precisely so that

$$\sum_{k \in K_m} f_k z^k = 0 \text{ for all } z \in I \text{ implies } f_k = 0 \text{ for all } k \in K_m.$$

For the sake of brevity, we only consider the case $n \geq 2$.

Theorem 6.2. *Suppose $n \geq 2$ and that the functions g_k are linearly independent and satisfy equation* (20)*. Then we must have $R \in F_2$, $b_i \in F_3$ for all $1 \leq i \leq d$ and $a_{ij} \in F_4$ for all $1 \leq i, j \leq d$. Furthermore, the coefficients are constrained in such a way that $A_k \in F_n$ for all k such that k such that $n - 1 \leq k_1 + \ldots + k_d \leq n$.*

The proof follows the same pattern as the proof of Theorem 2.1 with heavier notation to account for the extra dimensions. We include it here for completeness.

Proof. First we show that the functions $A_k \in F_n$ are polynomials for all $k \in K_n$. Let $N = \binom{n+d}{n}$ be the cardinality of set K_n. Since the functions g_k are linearly independent, we can find N distinct points $x_1, \ldots, x_N$ independent of z such that the matrix with i-th column formed by vector $(g_k(x_i), k \in K_n)$ is non-singular. Now fix any z, we can rewrite condition (20) as a set of N simultaneous linear equations with N unknowns $A_k(z)$. Therefore the

solution exists and is unique and can be written as linear combinations of the monomials z^k, hence all of the $A_k(z)$ are polynomials in d variables of total degree less or equal to n.

In what follows, we will find it convenient to introduce the notation

$$(a)_i = (0, \dots, 0, a, 0, \dots, 0) \quad \text{where } a \text{ is the } i\text{-th component.}$$
$$(a,b)_{i,j} = (0, \dots, a, \dots, b, \dots, 0) \quad \text{where } a \text{ is the } i\text{-th component and } b \text{ is the } j\text{-th component.}$$

for certain $\mathbb{Z}_+^d$-valued indices, where $a, b \in \mathbb{Z}_+$.

Since we must have $A_k(z) \in F_n$ for all $k \in K_n$, we can conclude for any $1 \le i, j \le d$

$$\begin{aligned} A_0(z) &= -R(z) && \in F_n \\ A_{(1)_i}(z) &= b_i(z) - z_i R(z) && \in F_n \\ A_{(1,1)_{i,j}}(z) &= b_i(z) z_j + b_j(z) z_i + a_{ij}(z) - z_i z_j R(z) && \in F_n \end{aligned}$$

Therefore we may conclude immediately that the functions $R(z), b_i, a_{ij}$ are polynomials. On the other hand

$$A_{(n)_i}(z) = n z_i^{n-1} b_i(z) + \frac{n(n-1)}{2} z_i^{n-2} a_{ii}(z) - z_i^n R(z) \in F_n$$

by cancelling the z_i^{n-2} factor, we may deduce that:

$$n z_i b_i(z) + \frac{n(n-1)}{2} a_{ii}(z) - z_i^2 R(z) \in F_2 \tag{21}$$

Similarly by considering $A_{(n-1)_i}, A_{(n-2)_i}, A_{(n-1,1)_{i,j}}$, we get

$$(n-1) z_i b_i(z) + \frac{(n-2)(n-1)}{2} a_{ii}(z) - z_i^2 R(z) \in F_3 \tag{22}$$

$$(n-2) z_i b_i(z) + \frac{(n-2)(n-3)}{2} a_{ii}(z) - z_i^2 R(z) \in F_4 \tag{23}$$

$$(n-1) z_i z_j b_i(z) + z_i^2 b_j(z) + \frac{(n-2)(n-1)}{2} z_j a_{ii}(z) + (n-1) z_i a_{ij}(z) - z_i^2 z_j R(z) \in F_3 \tag{24}$$

Subtracting (21) from (22) and subtracting (22) from (23) gives

$$\begin{aligned} z_i b_i(z) + (n-1) a_{ii}(z) &\in F_3 \\ z_i b_i(z) + (n-2) a_{ii}(z) &\in F_4 \end{aligned}$$

Hence we get the required degree constraint on functions R, b, a. For the remaining part of this theorem, observe that given the degree constraint, the functions A_k will automatically be in F_n as long as $k_1 + \ldots + k_d \le n-2$. □

7. Acknowledgement

The SC acknowledges the financial support of the Man Group studentship and MRT that of the Cambridge Endowment for Research in Finance. This work has been presented at the Labex Louis Bachelier - SIAM-SMAI Conference on Financial Mathematics in Paris, the Conference on Stochastic Calculus, Martingales and Financial Modeling in St Petersburg, the Advanced Methods in Mathematical Finance in Angers, and the Eleventh Cambridge–Princeton Conference in Princeton. We would like to thank the participants for useful

comments and suggestions. Finally, we would like to thank Dorje Brody and Lane Hughston for a discussion of this work and its relation to their paper [1].

References

[1] D.C. Brody and L.P. Hughston. Chaos and coherence: a new framework for interest-rate modelling. *Proceedings of the Royal Society of London A* **460**: 85-110. (2004)

[2] L. Chen, D. Filipović and H.V. Poor. Quadratic term structure models for risk-free and defaultable rates. *Mathematical Finance* **14**(4): 515-536 (2004)

[3] Cheng, Si. *Polynomial Models and Extensions in Mathematical Finance.* PhD thesis, University of Cambridge. (2016)

[4] J.C. Cox, J.E. Ingersoll and S.A. Ross. A theory of the term structure of interest rates. *Econometrica* **53**: 385-407 (1985)

[5] Ch. Cuchiero, M. Keller-Ressel, and J. Teichmann. *Polynomial processes and their applications to mathematical finance. Finance and Stochastics* **16**: 711-740 (2012)

[6] D. Duffie, D. Filipović, and W. Schachermayer. Affine processes and applications in finance. *Annals of Applied Probability* **13**(3): 984-1053 (2003)

[7] D. Duffie and R. Kan. A yield-factor model of interest rates. *Mathematical Finance* **6**(4): 379–406. (1996)

[8] Federal Reserve Bank of St. Louis. *Federal Reserve Economic Data.* https://fred.stlouisfed.org/ accessed August 2016.

[9] D. Filipović. Separable term structures and the maximal degree problem. *Mathematical Finance* **12**(4): 341-349 (2002)

[10] D. Filipović and M. Larsson. Polynomial preserving diffusions and applications in finance. *Finance and Stochastics.* doi:10.1007/s00780-016-0304-4 (2016)

[11] V. Goodman. Brownian super-exponents. arXiv:math/0612160 (2006)

[12] F. Jamshidian. Bond, futures and option evaluation in the quadratic interest rate model. *Applied Mathematical Finance* **3**: 93-115 (1996)

[13] I. Karatzas and S.E. Shreve. *Brownian motion and stochastic calculus.* Second edition. Graduate texts in mathematics. Springer 1998.

[14] M. Leippold and L. Wu. Asset pricing under the quadratic class. *Journal of Financial and Quantitative Analysis* **37**(2): 271-295 (2002)

[15] F.A. Longstaff. A nonlinear general equilibrium model of the term structure of interest rates. *Journal of Financial Economics* **23**: 195-224 (1989)

[16] A.F. Siegel. Price-admissibility conditions for arbitrage-free linear price function models for the term structure of interest rates. *Mathematical Finance* doi: 10.1111/mafi.12075 (2014)

[17] O. Vasicek. An equilibrium characterisation of the term structure. *Journal of Financial Economics* **5**(2): 177-188. (1977)

Statistical Laboratory, Centre for Mathematical Sciences, Wilberforce Road, Cambridge CB3 0WB, UK

E-mail address: `sc591@cam.ac.uk`, `m.tehranchi@statslab.cam.ac.uk`